

## Evaluation of a Reference Material for Glycated Haemoglobin

Cas W. Weykamp<sup>1</sup>, Theo J. Penders<sup>1</sup>, Frits A. J. Muskiet<sup>2</sup> and Willem van der Slik<sup>2</sup>

<sup>1</sup> Department of Clinical Chemistry, Queen Beatrix Hospital, Winterswijk, The Netherlands

<sup>2</sup> Central Laboratory for Clinical Chemistry, University Hospital Groningen, Groningen, The Netherlands

**Summary:** The use of lyophilized blood as a reference material for glycated haemoglobin was investigated with respect to IFCC criteria for calibrators and control materials. Ninety-two laboratories, using 11 methods, detected no changes in glycated haemoglobin content when the lyophilizate was stored for one year at 4 °C. Affinity chromatography, HPLC, electrophoresis and immunoassay detected no changes following 18 months storage at –84 and –20 °C. Samples for HPLC are stable at 4 °C for one year, and 5 years at –20 °C. For the other three methods, samples are stable for 5 years at 4 °C. At 4 °C, reconstituted samples are stable for 2 days (HPLC) and 7 days (other three). Lyophilization does not cause matrix effects and inhomogeneity, since mean glycated haemoglobin and reproducibility for lyophilized samples and whole blood were similar. The coefficient of variation for vial filling precision was 0.59%. We conclude that lyophilized blood samples can be used as calibrators and control materials. Their use as calibrators, following assignment of the HbA<sub>1c</sub> value by HPLC, may contribute, in the interim, to the standardized interpretation of long term diabetic control.

### Introduction

Glycated haemoglobin is a valuable indicator for long-term diabetic control (1). At present, mean intra-laboratory CVs amount to 5.2% (range: 0.2–28.7%), whereas inter-laboratory CV is about 20% (24% and 20% at glycated haemoglobin values of 5.0% and 15.8%, respectively) (2, 3). It implies that few laboratories meet the recommended intra-laboratory CV of < 5% (4), or the clinically more desirable CV of < 3.3% (5), and that glycated haemoglobin results from different laboratories are hardly comparable. Measurement of different analytes (e.g. glycated haemoglobin or HbA<sub>1c</sub>), methodological imperfections and lack of a primary standard and reference method contribute to the present dispersion of glycated haemoglobin results (6, 7). Standardization by the use of calibrators with assigned values has been proposed (8). In a recent pilot study, we (9) showed that three-point calibration reduces intra-laboratory CV from 6.6% to 3.5%, and inter-laboratory CV from 25 to 7% (at 5.5% glycated haemoglobin and from 15 to 4% (at 14.1% glycated haemoglobin).

Calibrators for the glycated haemoglobin standardization study were prepared by lyophilization of dialysed and subsequently haemolysed EDTA-blood from non-diabetic and diabetic volunteers (3). This procedure has been used for many years for the distribution of quality assurance samples in an external quality control programme in The Netherlands (3). Suitability of the method for the preparation of reference material and calibrators according to IFCC recommendations for cali-

brators and control materials (10) was, however, not tested. In this study we investigated the stability of both lyophilized and reconstituted material, matrix effects and the homogeneity of the candidate reference material.

### Materials and Methods

#### Samples

Lyophilized samples, stored in vacuum sealed vials, were prepared by the SKZL (the Dutch Foundation for Quality Control in Hospital Laboratories) according to the previously described procedure (3).

#### Stability of lyophilized material

One batch of 630 vials was prepared in July 1991. The vials contained lyophilized reference material of a single haemolysate pool with a glycated haemoglobin of 9.8%, as determined by HPLC (Diamat Bio-Rad, Bio-Rad Laboratories, Brea, CA, USA).

Four sets of 120 vials each were immediately stored at 4 °C in the dark. They were successively distributed to serve in the regular external quality control programme. Distribution was by post at ambient temperatures. The participants were unaware of the similarity of the material. The 4 sets were analysed 2.5, 6, 9 and 12 months after preparation, respectively.

Immediately after preparation the remaining 150 vials were stored in 5 sets of 30 vials each at –84, –20, 4, 22, and 37 °C, respectively. Five vials stored at each of the above temperatures were used for glycated haemoglobin analyses after 0, 3, 6, 9, 12 and 18 months.

#### Stability of reconstituted material

Two sets of 30 vials, containing lyophilized material from EDTA-blood samples with low (5.5%; HPLC, Diamat Bio-Rad) and high

(11.3%) glycated haemoglobin, respectively, were prepared. On days 0 (day of lyophilization-completion), 2, 4, 5, 6 and 7 we reconstituted 5 samples of each set with water to give final glycated haemoglobin concentrations equal to those in EDTA-blood. Reconstituted samples were stored at 4 °C in the dark. Their glycated haemoglobin fractions were analysed in one series on day 7; i.e. after 0, 1, 2, 3, 5 and 7 days storage.

#### Matrix-effects

EDTA-blood samples were collected from a non-diabetic (glycated haemoglobin 5.3%; HPLC, Diamat Bio-Rad) and diabetic (11.4%) volunteer. After division of each sample in two aliquots, one part was immediately processed to lyophilized samples and the other stored at 4 °C for 2 days (i.e. until completion of the lyophilization process). Immediately after lyophilization was completed, glycated haemoglobin was analysed in both lyophilized material and stored EDTA-blood samples. Each sample was analysed 10-fold.

#### Homogeneity

Homogeneity was tested during the preparation of 1700 vials, each from haemolysates with low (5.3%; HPLC, Diamat Bio-Rad) and high (11.4%) glycated haemoglobin. At intervals of 100 vials a total of 17 vials was taken during the filling procedure. Empty and filled vials were weighed to determine filling precision. Separate sets of 10 vials each were randomly selected for single analyses of glycated haemoglobin. Results were compared with 10-fold analyses in original haemolysates in the same series. Original haemolysates had been stored at 4 °C for 2 days (i.e. until completion of the lyophilization process).

#### Analytical methods

For the establishment of stability, matrix effects and homogeneity we used 4 commercially available glycated haemoglobin tests, according to the manufacturers' instructions. The tests were: Pierce affinity chromatography (Pierce, Glycotest II, Pierce, Rockford, IL, USA); Bio-Rad Diamat HPLC (Bio-Rad Diamat, Bio Rad Laboratories, Brea, CA, USA); Beckman electrophoresis (Beckman, Diatrac, Beckman Instruments, Brea, CA, USA); and the DAKO immunochemical method (ITK, HbA<sub>1c</sub>, DAKO, Cambridge, UK).

For stability studies in the external quality control programme, the participants also used the following commercially available methods, all but one according to the manufacturers' instructions: modified Pierce affinity chromatography (modification: 10 ml instead of 5 ml elution wash buffer; Glycotest II, Pierce, Rockford, IL, USA); Isolab affinity chromatography (Glyc-Affin; Isolab, Akron, OH, USA); HPLC Bio Rad Modular (Bio-Rad Laboratories and designed for HbA<sub>1c</sub> with Bio-Rad TSK column, Brea, CA, USA); Pharmacia HPLC (various HPLC systems equipment with Pharmacia Mono-S HR 5/5 columns; Pharmacia-LKB Technology, Uppsala, Sweden); Disposable ion exchange columns (HbA<sub>1c</sub> test; Bio-Rad Laboratories, Brea, CA, USA); Corning electrophoresis (Corning Medical Ltd, Halstead, UK); and Sebia electrophoresis (Hydrated glycosylated haemoglobins, Sebia, Issy-les-Moulineaux, France).

#### Statistics

Statistically significant differences between mean glycated haemoglobin values and SDs were determined with *Student's t*- and *F*-tests (9), respectively, at  $p < 0.05$ . We used two-way ANOVA to compare the combined effects of different laboratories and at different times (tab. 1). Repeated measurements were used to evaluate the glycated haemoglobin changes during storage-experiments shown in tables 2 and 3 (12).

### Results

#### Stability of lyophilized material

Table 1 shows mean glycated haemoglobin fractions and interlaboratory SDs, as determined by 92 laboratories

using 11 different methods. Lyophilized samples were analysed after 2.5, 6, 9 and 12 months storage at 4 °C. Some methods were performed by only a few laboratories. Their results should therefore be interpreted with caution. Except for results for the sample that was stored for 9 months and analysed by affinity chromatography (Pierce-Miedema), none of the methods showed significant changes in glycated haemoglobin fractions during 12 months storage (according to ANOVA). The dispersion of glycated haemoglobin fractions measured in different laboratories using the same method (expressed as the interlaboratory SD) was invariable.

Table 2 shows selected results of the experiment with lyophilized samples that had been stored for various months at various temperatures. Immediately after lyophilization, the fraction of glycated haemoglobin in the batch was determined with 4 methods (affinity chromatography, HPLC, electrophoresis, immunoassay) in a single laboratory. Vials were stored at -84, -20, 4, 22 and 37 °C, respectively. Glycated haemoglobin was re-analysed after storage for 3, 6, 9, 12 and 18 months. Table 2 shows the initial glycated haemoglobin fractions and the fractions found after 18 months storage. Glycated haemoglobin fractions determined by HPLC in samples that had been stored at -84 and -20 °C were not different from initial glycated haemoglobin fraction. The values for glycated haemoglobin in samples stored at 4, 22 and 37 °C were, however, lower than the initial glycated haemoglobin and lower than the value for glycated haemoglobin in the sample stored at -84 °C. Figure 1 shows HPLC profiles of lyophilized samples stored for 18 months at -84 (a) and 37 °C (b), respectively. The sample stored at 37 °C contained an additional peak, that influenced both HbA<sub>1c</sub> and HbA<sub>0</sub> integration. Glycated haemoglobin values determined by electrophoresis and affinity chromatography were lower in all stored samples, compared with initial values. No significant differences were observed between samples stored at -84, -20, 4 and 22 °C. Compared with the sample that was stored at -84 °C, the sample stored at 37 °C showed a higher glycated haemoglobin fraction in affinity chromatography, and a lower glycated haemoglobin in electrophoresis. As determined by immunoassay, the glycated haemoglobin of all stored samples was higher than the initial glycated haemoglobin, except in the case of the sample stored at -84 °C, which showed no difference. Results of 3, 6, 9 and 12 months storage are not shown. For storage at 37 °C the first deviating results (according to repeated measurements) were observed at 3 months (HPLC; lower glycated haemoglobin), 9 months (affinity chromatography (higher glycated haemoglobin) and 12 months (electrophoresis; lower glycated haemoglobin). At 22 °C, deviating results were only observed with HPLC at 6 months (lower glycated haemoglobin).

### Stability of reconstituted material

Lyophilized samples were reconstituted on days 0, 2, 4, 5, 6 and 7, stored in the interim at 4 °C, and analysed on day 7 in one series. Table 3 shows glycated haemoglobin results for the freshly reconstituted sample (storage time 0 days), together with results from samples that were reconstituted 1, 2, 3, 5 and 7 days before analysis, respectively. The data were obtained with 4 methods and applied to samples that contained low and high glycated haemoglobin fractions, respectively. Repeated measurements with affinity chromatography, electrophoresis and immunoassay revealed no changes in the glycated haemoglobin values. When measurements were performed

with HPLC, glycated haemoglobin was found to be decreased in samples that had been stored for 3 days or longer (low glycated haemoglobin) and 5 days or longer (high glycated haemoglobin).

### Matrix effects

To investigate possible matrix effects introduced by the lyophilization procedure, glycated haemoglobin fractions were determined in a whole blood sample that had been stored for 2 days at 4 °C, and in the same sample that been lyophilized in that period. Results of analyses of samples with low and high glycated haemoglobin fractions by 4 methods, are shown in table 4. Lyophiliza-

**Tab. 1** Mean  $\pm$  SD glycated haemoglobin fractions<sup>a</sup> per method as determined by 92 laboratories in a lyophilized sample stored at 4 °C for 2.5, 6, 9 and 12 months.

Method	No. of laboratories	Storage at 4 °C			
		2.5 months	6 months	9 months	12 months
Affinity Chromatography					
Pierce-Miedema	34	8.4 ± 0.7	8.2 ± 0.8	8.0 <sup>b</sup> ± 0.6	8.2 ± 0.8
Pierce	3	10.0 ± 0.2	10.7 ± 0.8	10.2 ± 0.5	9.7 ± 0.3
Isolab	1	10.8	10.0	11.6	11.2
HPLC					
Bio Rad Diamat	7	10.1 ± 0.6	9.7 ± 0.3	10.0 ± 0.3	9.9 ± 0.7
Bio Rad Modular	13	9.8 ± 0.6	9.9 ± 0.4	9.7 ± 0.6	9.6 ± 0.4
Pharmacia	7	9.0 ± 0.6	9.3 ± 0.8	9.0 ± 1.1	9.2 ± 1.1
Electrophoresis					
Beckman	15	9.2 ± 1.3	9.5 ± 1.1	9.3 ± 1.2	9.2 ± 0.1
Corning	3	11.1 ± 0.7	10.3 ± 0.9	10.8 ± 0.8	11.2 ± 0.7
Sebia	1	12.0	11.7	11.9	11.9
Disposable ion exchange					
Bio Rad HbA1c	2	9.3 ± 0.8	8.9 ± 0.4	10.0 ± 0.6	9.9 ± 0.5
Immuno Assay					
DAKO	1	7.4	7.2	7.0	7.4

<sup>a</sup> Mean  $\pm$  inter-laboratory SD

<sup>b</sup> Significantly different from mean glycated haemoglobin after 2.5 months ( $p < 0.05$ ).

**Tab. 2** Mean  $\pm$  SD glycated haemoglobin fractions<sup>a</sup> as determined by 4 methods in a lyophilized sample stored at various temperatures during 18 months.

Storage temperature in °C	Affinity chromatography Pierce	HPLC Bio Rad – Diamat	Electrophoresis Beckman	Immuno assay DAKO
<i>Initial fractions of glycated haemoglobin<sup>b</sup></i>				
	9.70 $\pm$ 0.18	9.80 $\pm$ 0.03	10.32 $\pm$ 0.47	7.27 $\pm$ 0.27
<i>After storage</i>				
–84	8.42 $\pm$ 0.20	9.81 $\pm$ 0.03	9.07 $\pm$ 0.53	8.23 $\pm$ 0.23
–20	8.53 $\pm$ 0.17	9.82 $\pm$ 0.03	9.31 $\pm$ 0.58	8.06 $\pm$ 0.23
4	8.28 $\pm$ 0.27	9.48 <sup>c</sup> $\pm$ 0.05	9.08 $\pm$ 0.43	8.04 $\pm$ 0.18
22	8.27 $\pm$ 0.23	8.83 <sup>c</sup> $\pm$ 0.04	9.52 $\pm$ 0.73	8.31 $\pm$ 0.26
37	9.35 <sup>c</sup> $\pm$ 0.38	8.56 <sup>c</sup> $\pm$ 0.17 <sup>d</sup>	8.03 <sup>c</sup> $\pm$ 0.66	8.09 $\pm$ 0.22

<sup>a</sup> Mean  $\pm$  SD for 5 assays

<sup>b</sup> Glycated haemoglobin fraction as determined in the sample before storage.

<sup>c</sup> Mean significantly different from mean of sample stored at –84 °C ( $p < 0.05$ ).

<sup>d</sup> SD significantly different from SD of sample stored at –84 °C ( $p < 0.05$ ).

tion did not affect the glycated haemoglobin results as obtained by affinity chromatography, electrophoresis and immunoassay. Analyses by HPLC showed small, but significant, glycated haemoglobin increase in the samples containing low (absolute increase 0.14%) and high (absolute increase 0.17%) concentrations of glycated haemoglobin. Figure 2 shows HPLC and electrophoretic profiles of whole blood and lyophilized blood, using the sample containing a high concentration of glycated haemoglobin. Lyophilization caused a small

peak (retention time 1.2 min) to disappear from the HPLC-chromatogram. No other differences were detected.

### Homogeneity

Mean filling weight amounted to 123.9 mg with a CV of 0.59%. The reproducibility of assays in whole blood and lyophilized blood is shown in table 4. Except for the analysis of the low glycated haemoglobin concentra-

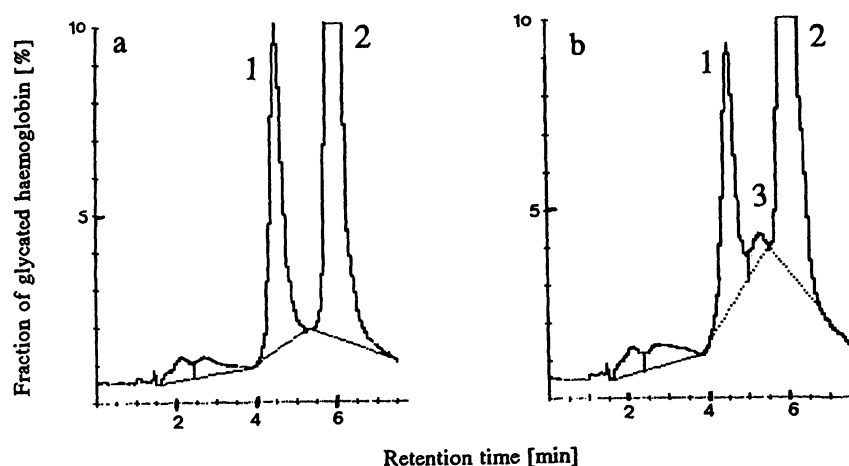


Fig. 1 Glycated haemoglobin profiles of lyophilized samples, as determined by HPLC (Diamat, Bio-Rad) after 18 months storage at  $-84^{\circ}\text{C}$  (a) and  $37^{\circ}\text{C}$  (b).

Peaks: 1 = HbA<sub>1c</sub>; 2 = HbA<sub>0</sub>, 3 = Decomposition product.

Tab. 3 Mean  $\pm$  SD glycated haemoglobin fractions<sup>a</sup> as determined by 4 methods in a reconstituted sample stored at  $4^{\circ}\text{C}$  during 0–7 days.

Storage time in days	Affinity chromatography Pierce		HPLC Bio Rad, Diamat		Electrophoresis Beckman		Immuno assay DAKO	
	Glycated haemoglobin		Glycated haemoglobin		Glycated haemoglobin		Glycated haemoglobin	
	low	high	low	high	low	high	low	high
0	$3.86 \pm 0.15$	$11.59 \pm 0.37$	$5.50 \pm 0.05$	$11.27 \pm 0.08$	$4.61 \pm 0.67$	$11.20 \pm 0.33$	$3.47 \pm 0.06$	$11.72 \pm 0.30$
1	$3.81 \pm 0.12$	$11.52 \pm 0.23$	$5.51 \pm 0.04$	$11.25 \pm 0.09$	$4.37 \pm 0.47$	$11.03 \pm 0.78$	$3.44 \pm 0.09$	$11.71 \pm 0.26$
2	$3.83 \pm 0.24$	$11.48 \pm 0.32$	$5.46 \pm 0.07$	$11.32 \pm 0.07$	$4.43 \pm 0.61$	$11.15 \pm 0.63$	$3.46 \pm 0.05$	$11.75 \pm 0.25$
3	$3.88 \pm 0.10$	$11.53 \pm 0.17$	$5.21^b \pm 0.11$	$11.22 \pm 0.12$	$4.71 \pm 0.38$	$11.29 \pm 0.48$	$3.48 \pm 0.08$	$11.68 \pm 0.17$
5	$3.87 \pm 0.17$	$11.62 \pm 0.43$	$4.94^b \pm 0.12^c$	$10.98^b \pm 0.15$	$4.68 \pm 0.71$	$11.38 \pm 0.73$	$3.43 \pm 0.03$	$11.72 \pm 0.41$
7	$3.84 \pm 0.17$	$11.55 \pm 0.35$	$4.82^b \pm 0.20^c$	$10.28^b \pm 0.24$	$4.52 \pm 0.58$	$11.16 \pm 0.63$	$3.43 \pm 0.09$	$11.73 \pm 0.32$

<sup>a</sup> Mean  $\pm$  SD for 5 assays.

<sup>b</sup> Mean significant different ( $p < 0.05$ ) from mean of freshly reconstituted sample (day 0).

<sup>c</sup> SD significantly different ( $p < 0.05$ ) from SD of freshly reconstituted sample (day 0).

Tab. 4 Mean  $\pm$  SD glycated haemoglobin fractions in whole blood and lyophilized samples with low and high glycated haemoglobin fractions, as determined by 4 methods.

Method	Low glycated haemoglobin		High glycated haemoglobin	
	Whole blood	After lyophilization	Whole blood	After lyophilization
Affinity chromatography (Pierce)	$3.77 \pm 0.17$	$3.81 \pm 0.14$	$11.61 \pm 0.24$	$11.74 \pm 0.30$
HPLC (Bio-Rad, Diamat)	$5.28 \pm 0.04$	$5.42^a \pm 0.04$	$11.37 \pm 0.06$	$11.54 \pm 0.09$
Electrophoresis (Beckman)	$4.24 \pm 0.31$	$4.32 \pm 0.65^b$	$10.97 \pm 0.49$	$10.76 \pm 0.38$
Immuno assay (DAKO)	$3.35 \pm 0.08$	$3.33 \pm 0.07$	$11.63 \pm 0.27$	$11.66 \pm 0.24$

<sup>a</sup> Mean after lyophilization significantly different from mean of whole blood ( $p < 0.05$ ).

<sup>b</sup> SD after lyophilization significant different from SD of whole blood ( $p < 0.05$ ).

tion by electrophoresis, none of the intra-sample SDs in the lyophilized haemolysates were found to be significantly different from those of the corresponding whole blood.

## Discussion

The IFCC recommendations for calibrators and control materials refer to stability, matrix effects and homogeneity as important quality criteria (10, 13). We tested these for the glycated haemoglobin reference material that is currently used in the SKZL external quality assessment programme.

We used two approaches to test the stability of the lyophilized material. In a multi-centre study 92 laboratories, using 11 methods, detected no differences in the glycated haemoglobin value over a period of one year. The second approach to test glycated haemoglobin stability was based on the assumption that (bio)chemical processes proceed 2–3 times faster with every 10 °C temperature increase (14). A reference material that proves stable for  $n$  months at  $x$  °C is expected to be stable for  $4n$  months at  $(x-20)$  °C. Samples stored at various temperatures were stable at the conventional refrigerator temperature of 4 °C for at least 5 years when analysed

by affinity chromatography, electrophoresis and immunoassay and for 1 year for HPLC analysis. The underlying cause of decreased stability with HPLC is the appearance of an extra peak that elutes between HbA<sub>1c</sub> and HbA<sub>0</sub>, and thereby affects integration properties (fig. 1). Formation of an interfering component has previously been noted and is often referred to as formation of HbA<sub>3</sub> (16). From both multi-centre and storage studies we conclude that the lyophilized material is essentially stable. Stability of the reconstituted material was investigated by storage at 4 °C during 0–7 days.

These storage-tests of the reconstituted material at 4 °C revealed 2 days stability for HPLC and 7 days for the other techniques. Again, HbA<sub>3</sub> formation limits the apparent stability in HPLC. Lyophilization and addition of cryoprotectants change the biological matrix and may thereby cause deviant glycated haemoglobin behaviour, compared with whole blood. Results show that lyophilization does not affect the concentration of glycated haemoglobin as determined by affinity chromatography, HPLC, electrophoresis and immunoassay. The comparable precision of glycated haemoglobin assays of lyophilized samples and whole blood indicates that the homogeneity of the lyophilized material is acceptable.

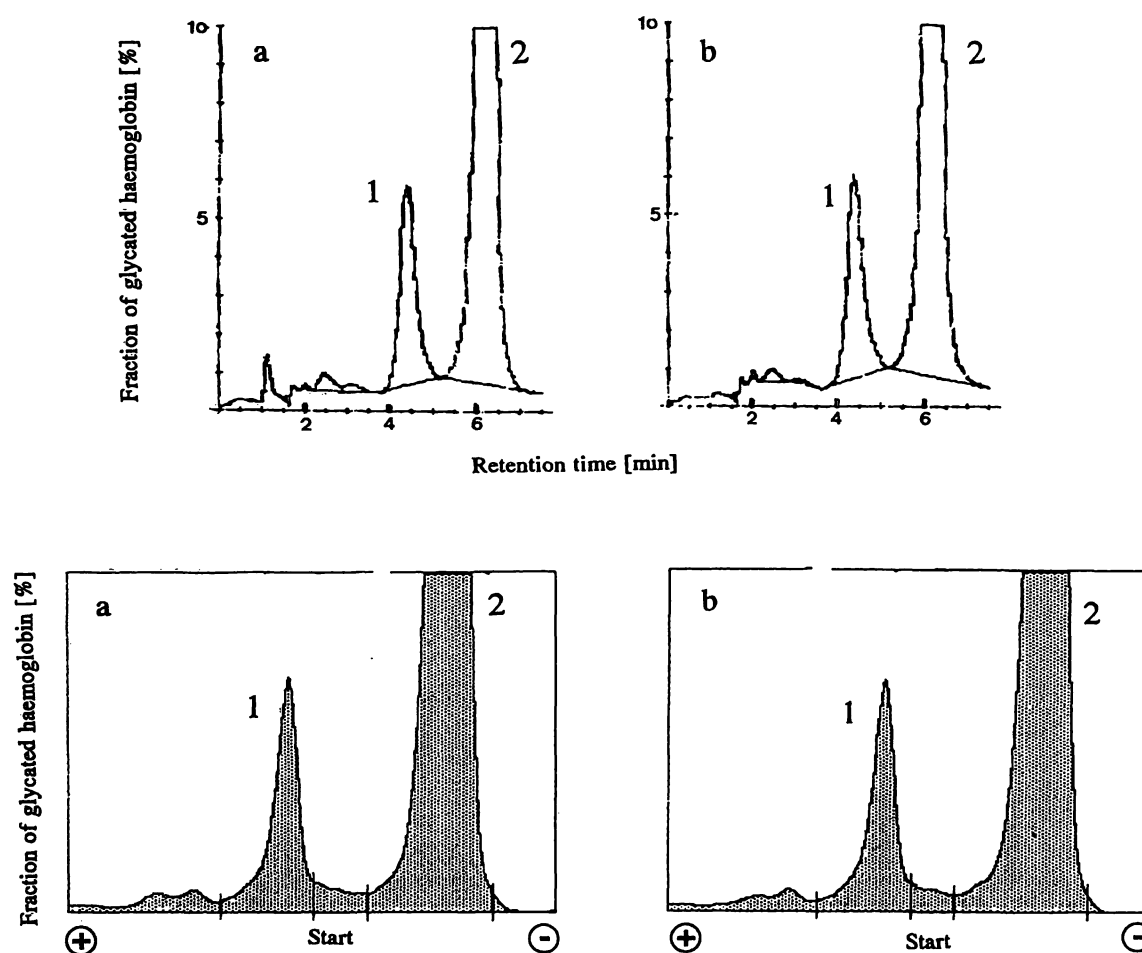


Fig. 2 Glycated haemoglobin profiles as determined by HPLC (Diamat, Bio-Rad; top-panels) and electrophoresis (Beckman; bot-

tom panels) for whole blood (a) and after lyophilization (b). S is the application point (a).

Peaks: 1 = HbA<sub>1c</sub>; 2 = HbA<sub>0</sub>.

Standardization of glycated haemoglobin assays is a prerequisite for a standardized interpretation of long term diabetic control. The need for standardized interpretation has been shown in the DCCT study (17). Since a basic solution is not to be expected before the end of this century, standardization by calibration (8, 9) is proposed as a pragmatic, interim, approach. This requires the use of a reliable calibrator, whose preparation is complicated by the brittle structure of haemoglobin. We conclude that the lyophilized material used in the SKZL external quality programme meets IFCC-criteria for stability, homogeneity and matrix effects. The samples can be used both as calibrators and control materials. Its acceptance for calibration requires reliable assignment of a

glycated haemoglobin value. Assignment of an HbA<sub>1c</sub> value on the basis of an absolute HbA<sub>1c</sub> method that uses a primary standard mixture of purified HbA<sub>1c</sub> and HbA<sub>0</sub> would be ideal. However, neither pure HbA<sub>1c</sub> nor an absolute method are available (15). Assignment of a glycated haemoglobin value on the basis of the HPLC method that was used in the DCCT study seems a reasonable interim solution for achieving a standardized interpretation of long term diabetic control.

### Acknowledgements

We thank Dr. H. Baadenhuijsen for performing the HPLC analyses of glycated haemoglobin.

### References

- Goldstein DE, Parker KM, England JD. Clinical application of glycosylated hemoglobin measurements. *Diabetes* 1982; 21 (3 Suppl):70-8.
- John WG. Performance of glycated hemoglobin analysis: data from the UK national external quality assessment scheme. *Clin Chem* 1990; 36:1002.
- Weykamp CW, Penders TJ, Muskiet FAJ, van der Slik W. Glycohaemoglobin: comparison of 12 analytical methods, applied to lyophilized haemolysates by 101 laboratories in an external quality assurance program. *Ann Clin Biochem* 1993; 30:169-74.
- Baynes J, Bunn H, Goldstein D. National diabetes data group: report of the expert committee on glycosylated hemoglobin. *Diabet Care* 1984; 7:605-15.
- Larsen ML, Fraser CG, Petersen PH. A comparison of analytical goals for haemoglobin A<sub>1c</sub> assays derived using different strategies. *Ann Clin Biochem* 1991; 28:272-8.
- Bruns DE. Standardization, calibration, and the care of diabetic patients. *Clin Chem* 1992; 38:2263-4.
- Weykamp CW, Penders TJ, Muskiet FAJ, van der Silk W. Influence of hemoglobin variants and derivatives on glycohemoglobin determinations, as investigated by 102 laboratories using 16 methods. *Clin Chem* 1993; 39:1717-23.
- Little RR, Wiedmeyer HM, England JD, Wilke AL, Rohlfing CL, Wians FH, et al. Interlaboratory standardization of measurements of glycohemoglobins. *Clin Chem* 1992; 38:2472-8.
- Weykamp CW, Penders TJ, Muskiet FAJ, van der Silk W. Effect of calibration on dispersion of glycohemoglobin values as determined by 111 laboratories using 21 methods. *Clin Chem* 1994; 40:138-44.
- International Federation of Clinical Chemistry, Scientific Committee, Expert Panel on Nomenclature and Principles of Quality Control in Clinical Chemistry. Approved recommendation (1979) on quality control in clinical chemistry. Part 3. Calibration and control materials. *J Clin Chem Clin Biochem* 1980; 18:855-60.
- Skoog DA, West DM. *Fundamentals of analytical chemistry*. 2nd ed. London: Holt, Rinehart and Winston, 1970:25.
- Snedecor GW, Cochran WG. *Statistical methods*. Ames, Iowa: Iowa State Univ. Press, 1978:258-338.
- Dybkaer R. Reference materials - a main element in a coherent reference measurement system. *Eur J Clin Chem Clin Biochem* 1991; 29:241-6.
- Jenne NK, Perry WLM. The stability of biological standards. *Bull Wld Hlth Org* 1956; 14:167-82.
- Goldstein DE, Little RR, Wiedmeyer HM, England JD, McKenzie EM. Glycated hemoglobin: methodologies and clinical applications. *Clin Chem* 1986; 32 (Suppl): B64-B70.
- Little RR, Wiedmeyer HM, England JD, Naito HK, Goldstein DE. Interlaboratory comparison of glycohemoglobin results: College of American Pathologists survey data. *Clin Chem* 1991; 37:1725-9.
- The Diabetes Control and Complications Trial Research Group. The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus. *N Eng J Med* 1993; 329:977-86.

Received June 20/September 18, 1995

Corresponding author: Cas W. Weykamp, Department of Clinical Chemistry, Queen Beatrix Hospital, Beatrixpark 1, NL-7101 BN Winterswijk, The Netherlands

# Evaluation of a Second-Generation Thyrotropin Automated Immunoassay

Laurent Boon-Falleur and Philippe De Nayer

Laboratoire de Médecine Nucléaire, Cliniques Universitaires Saint-Luc, Université Catholique de Louvain, Bruxelles, Belgium

**Summary:** We evaluated the analytical performance of the Vista automated immunoassay system for human thyrotropin determination. The operating characteristics as well as the analytical performance were assessed. The Vista human thyrotropin immunoassay showed a minimal detection limit of 0.08 mU/l and a functional sensitivity of 0.12 mU/l. The system meets the criteria for second-generation human thyrotropin assays.

## Introduction

Immunometric assays for thyrotropin have replaced the less sensitive thyrotropin radioimmunoassays. Immunometric assays have lower detection limits and possibly a faster turnaround time. Second and third-generation human thyrotropin assays (20% inter-assay CV at 0.1 to 0.2 mU/l and 0.01 to 0.02 mU/l) are becoming available on automated immunoassay analysers (1, 2). The transition between radiometric techniques and fluorometric or chemiluminescent techniques requires careful assessment of the performance of the newer kits. Indeed, the use of a non-isotopic signal is not, per se, a guarantee of quality.

This report describes the evaluation of an automated immunoassay (Vista hTSH, Syva) for detection of human thyrotropin. This kit is compared with an IRMA method (Riabead II, ABBOTT).

## Materials and Methods

### Sera and serum pools

Samples were collected in dry tubes, separated, and the serum frozen at  $-20^{\circ}\text{C}$ . For intra- and inter-assay purposes, routine samples were pooled according to their human thyrotropin values and aliquoted before freezing at  $-20^{\circ}\text{C}$ .

### Controls

We used human serum pools containing different concentrations of human thyrotropin to evaluate the intra- and inter-assay precision of the system. Routine sera were used to correlate the results of this system with the Abbott method.

### Method used for comparison

Immunoradiometric assay (IRMA): Riabead II (Abbott). This assay shows a 20% inter-assay CV at 0.1 mU/l.

### Method under evaluation

The Vista assay is a two-step sandwich method based on two monoclonal antibodies. The capture antibody, specific for the intact human thyrotropin molecule, is coupled to chromium dioxide para-

magnetic particles ( $\text{CrO}_2$ ) (3). This solid phase system presents a very large surface area, compared with more classical solid-phase techniques such as coated tubes, polystyrene beads or microtitre plates. The  $\text{CrO}_2$  particle is separated by a magnetic procedure. The second antibody, specific for the  $\beta$  subunit of human thyrotropin molecule, is conjugated to alkaline phosphatase.

### VISTA analyser

The Vista Immunoassay system is a benchtop multitest analyser that provides a large panel of methods (max. 15 in the same run). The system can handle 50 different samples and perform a maximum of 100 measurements in a run.

This analyser provides on-board refrigerated reagent storage, eliminating the need for daily manipulations of reagents before each run. Calibration curves are kept in memory for 30 days (human thyrotropin assay).

The sample (100  $\mu\text{l}$ ) is dispensed by the system. The  $\text{CrO}_2$  particles and the second antibody are then added and incubated for a period of 30 min at  $37^{\circ}\text{C}$ . The particles are washed and 4-methyl-umbelliferyl phosphate is added to produce the fluorescent signal.

### Calibration

The human thyrotropin assay was calibrated using the calibrators provided with the system (5 point calibration curve). The calibration curve for a reagent lot is stable for 30 days. Two calibration curves for two different reagent lots can coexist in memory. Before each run, the system automatically fits its calibration curve(s) for the different assays by a two point normalization of the curve(s), using a quinine sulphate solution to standardize the fluorometer.

## Results

### Operating characteristics

During our testing period, we found that a mean of 30 minutes was enough to start the system every morning, including the time for pipetting samples into the analyser cups. The FT4 assay was also evaluated during the same runs (unpublished data). Each run was complete after 140 minutes, but all results are available after 120 minutes. The remaining time is consumed by an extensive washing procedure for the reaction cuvettes, prior to their disposal (according to United States legal constraints). During the run, the analyser is truly 'walk-

away' in routine use. This time schedule easily allows the performance of three different runs in a day, i.e. 150 samples can be assayed and a maximum of 300 results can be obtained.

#### Intra-assay reproducibility

Ten pools covering the entire range of the human thyrotropin assay were aliquoted and measured 10 times in the same run. Table 1 summarizes the values obtained. A CV below 20% is obtained for human thyrotropin values above 0.05 mU/l (mean value of 0.05 mU/l: CV of 19.2%).

#### Inter-assay reproducibility

Ten pools, covering the entire range of the human thyrotropin assay were aliquoted, frozen at  $-20^{\circ}\text{C}$  and assayed in 10 different runs on different days. Table 1 presents the collected data. A CV below 20% is obtained for human thyrotropin values estimated at around 0.12 mU/l. This is commonly taken to represent the functional sensitivity of an assay (2).

#### Analytical sensitivity

To determine the analytical sensitivity limit of the assay, the zero standard was assayed ten times in the same run.

The mean fluorescence level (measured in Fluorescence Units, FU) + 2 SD (Standard Deviation: 6.2 FU) was

522 FU. This is lower than any fluorescence level observed for any sample assayed.

The concentration mean for standard 0 derived from the calibration curve is negative ( $-0.012$  mU/l; SD =  $0.006$  mU/l).

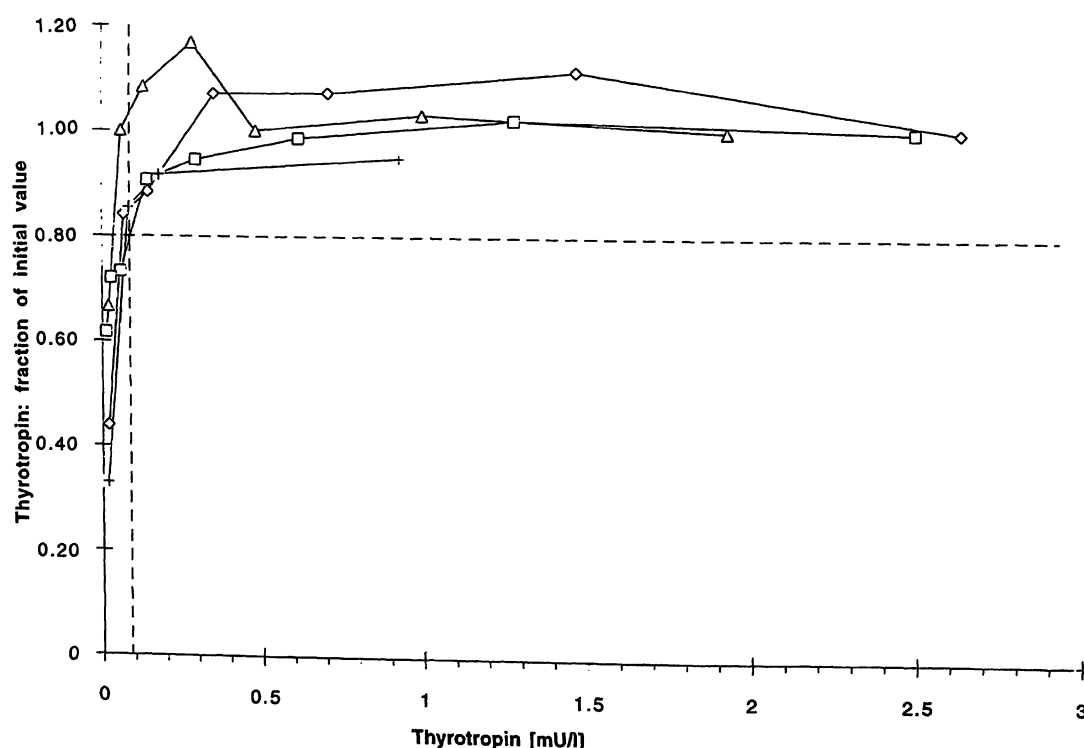
#### Minimal detection limit

The minimal detection limit is commonly evaluated by a dilution test. We selected single samples and per-

**Tab. 1** Intra- and inter-assay reproducibility.

The mean thyrotropin value obtained for ten replicates of the sera is tabulated for intra- and inter-assay against the CV obtained for those values. The precision cutoff is fixed at 20% CV. For intra-assay, this precision is obtained for thyrotropin values  $\geq 0.05$  mU/l. For inter-assay, this precision is obtained for thyrotropin values estimated above 0.12 mU/l.

Intra-assay		Inter-assay	
Mean thyrotropin (mU/l)	CV (%)	Mean thyrotropin (mU/l)	CV (%)
0.01	59.8	0.01	97.6
0.05	19.2	0.04	24.9
0.14	5.5	0.05	26.5
0.25	5.5	0.19	13.4
0.45	2.2	0.36	5.4
0.92	5.1	0.88	5.3
1.80	2.5	1.64	7.0
2.89	3.4	2.34	15.2
12.65	2.6	6.74	8.0
53.81	2.6	36.80	8.7



**Fig. 1** Minimal detection limit.

Serial dilution was performed for four samples. This figure presents the thyrotropin values of the diluted samples against the recovery, expressed in fraction of the initial value after correction for

dilution. The cutoff is fixed at the thyrotropin concentration for which the calculated recovery shows a 20% variation from the initial value. A cutoff value of 0.08 mU/l is obtained.



formed serial dilutions with the diluent provided with the kit. The dilutions were assayed in a single run for each sample. The cut-off was fixed at the calculated value (corrected for dilution) showing a 20% variation from the initial value. Figure 1 presents the curves of the obtained values. A minimal detection limit of 0.08 mU/l was obtained by this method.

#### Correlation with Abbott method

Samples ( $n = 118$ ) were assayed in the two methods (Abbott ( $x$ ) and Vista ( $y$ )). Over the entire range of human thyrotropin concentrations (from 0 to 40.6 mU/l), there was a good agreement between the two methods.

$$(y = 0.93x + 0.03 \quad r = 0.99)$$

For human thyrotropin concentrations less than 1 mU/l (fig. 2), the Vista method gave slightly lower values than the Abbott assay.

$$(y = 0.90x + 0 \quad r = 0.96)$$

#### Interference of human anti-mouse immunoglobulin antibodies

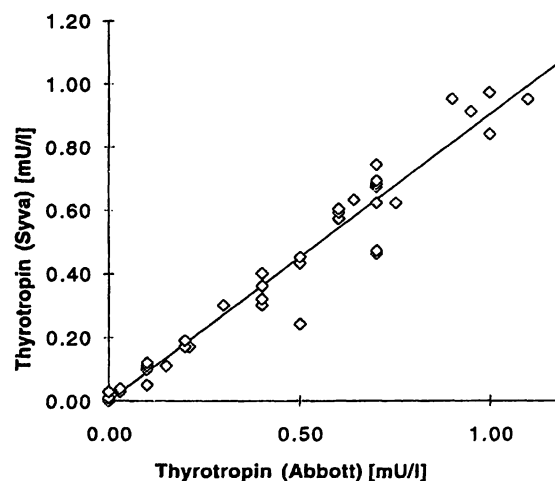
Three sera from a single patient treated with OKT3 (murine monoclonal antibody) were selected for this study. This patient developed human anti-mouse immunoglobulin antibodies after treatment (4).

Table 2 compares data from sera with and without addition of normal mouse serum, measured with the Abbott RIA method and the Vista assay. The Vista reagents clearly eliminate the interference which is seen for the Abbott method in the absence of normal mouse serum.

#### Discussion

Over the last ten years, the introduction of immunometric methods and non-radioisotopic labels have resulted in significant progress in the measurement of thyrotropin. The gain in sensitivity has led to a generational classification based on the minimal concentration yielding a between-assay CV of 20% (2). The latter figure is close to the minimum detection limit interpolated from the precision-dose profile at a CV of 22%, as proposed by McConway et al. in a study of ten manual IRMAs or non-isotopic immunometric assays for human thyrotropin (5).

The routine use of non-isotopic labels in automated assay systems is growing in clinical laboratories. Although second and third-generation sensitivity (inter-assay CV of 20% respectively at 0.1–0.2 mU/l and 0.01–0.02 mU/l) methods are available for manual operation, the transfer of this methodology to automated versions has yet to be validated.



**Fig. 2** Correlation between Syva Vista and Abbott Riabead II assay.

Patient sera were assayed by the Vista method and correlated with values obtained by the Abbott Riabead II method. The sera were selected for thyrotropin values  $\leq 1$  mU/l with the Vista method. ( $y = 0.90x + 0$ ,  $r = 0.96$ ,  $n = 44$ )

**Tab. 2** Interference of human anti-mouse immunoglobulins antibodies.

The sera were assayed by the Vista method and by the Abbott Riabead II method, with and without addition of 10  $\mu$ l of normal mouse serum.

Serum	Thyrotropin (mU/l)			
	Abbott without normal mouse serum	Abbott with normal mouse serum	Vista without normal mouse serum	Vista with normal mouse serum
A <sup>1</sup>	1.9	1.1	1.15	1.05
B <sup>2</sup>	33.3	1.7	0.99	— <sup>a</sup>
C <sup>2</sup>	46.0	1.1	1.10	1.03

<sup>1</sup> before treatment

<sup>2</sup> after treatment with OKT3 (murine monoclonal antibody)

<sup>a</sup> not enough serum to be assayed

A recent report indicates that for automated assays, third-generation functional sensitivity is not consistently reached in clinical practice despite the claim of manufacturers (6). These authors and others (7) also point to the need for improving the robustness of assays — even second-generation assays — to ensure their performance in clinical laboratory routine.

The results obtained in the present study with the VISTA immunoanalyser indicate that the performance of the thyrotropin assay corresponds to second-generation sensitivity. This IMA, based on a fluorescent label, appears to be equivalent to the Abbott IRMA technique used as the comparison method.

#### Acknowledgements

The assay material used for this study was kindly provided by the Syva company.

## References

1. Wilkinson E, Rae PWH, Thomson KJT, Toft A, Spencer CA, Beckett GJ. Chemiluminescent third-generation assay (Amerlite TSH-30) of thyroid-stimulating hormone in serum or plasma assessed. *Clin Chem* 1993; 39:2166-73.
2. Spencer CA, Lopresti JS, Patel A, Guttler RB, Eigen A, Shen D, et al. Applications of a new chemiluminometric thyrotropin assay to subnormal measurement. *J Clin Endocrinol Metab* 1990; 70:453-60.
3. BirkMeyer RC, Diaco R, Hutson DK, Lau HP, Miller WK, Neelkantan NV, et al. Application of novel chromium dioxide magnetic particles to immunoassay development. *Clin Chem* 1987; 33:1543-7.
4. Goffin E, Lambert M, DeNayer P, Saint-Remy JM, Pirson Y, van Ypersele de Strihou C. Spurious elevation of serum thyrotropin (TSH) after OKT3 administration. *Nephrol Dial Transplant* 1994; 9:1500-2.
5. McConway MG, Chapman RS, Beastall GH, Brown E, Tillman J, Bonar JA, et al. How sensitive are immunometric assays for thyrotropin? *Clin Chem* 1989; 35:289-91.
6. Spencer CA, Takeuchi M, Kazarosyan M, MacKenzie F, Beckett G, Wilkinson E. Interlaboratory/intermethod differences in functional sensitivity of immunometric assays of thyrotropin (TSH) and impact on reliability of measurement of subnormal concentrations of TSH. *Clin Chem* 1995; 41:367-74.
7. Pilo A, Zucchelli GC, Malvano R, Clerico A, Iervasi G, Signorini C. Interassay variability of immunometric methods for thyrotropin in an external quality assessment survey: evidence that functional sensitivity is not always adequate for clinical decisions. *Clin Chem* 1992; 38:1345-9.

*Received May 20/August 14, 1995*

Corresponding author: Dr. L. Boon-Falleur, Informatique des Laboratoires, Cliniques Universitaires Saint-Luc, av. Hippocrate, 10 / CTL-1730, B-1200 Bruxelles, Belgium

## CA 15-3: A Multicentre Evaluation of Automated and Manual Tests

Peter R. Huber<sup>1</sup>, Paul Bischof<sup>2</sup>, Robert Kretschmer<sup>3</sup>, Martie Truschnig<sup>4</sup>, Gabriele Halwachs<sup>4</sup> and Marianne Schmidt<sup>4</sup>

<sup>1</sup> Hormone Laboratory, Central Laboratory Division, University of Basle, Basle, Switzerland

<sup>2</sup> Department of Gynecology, University of Geneva, Geneva, Switzerland

<sup>3</sup> Department of Clinical Chemistry, University of Berne, Inselspital, Berne, Switzerland

<sup>4</sup> Medical University Hospital, Graz, Austria

**Summary:** The introduction of a new and automated CA 15-3 immunoassay (IMx Abbott) prompted us to compare the analytical performance of this new test with established tests from CIS ELSA, Sorin, and Boehringer Mannheim in a multicentre study.

CA 15-3 measurements in blood samples of breast tumour patients, comparison of intra- and inter-assay variation, dilution linearity, and lower limit of detection are described. The study showed improved precision for the automated over the manual test systems (intra-assay variation: IMx < 5%, CIS ELSA 4–9%, ES 300 < 3% and ETI Sorin > 10%; inter-assay variation: IMx ≤ 8%, CIS ELSA ≤ 19%, ES 300 ≤ 9% and ETI Sorin ≤ 27%).

Results on patients' samples (n = 101 to 184) showed highly comparable results; IMx vs CIS ELSA site 1: r = 0.950; IMx vs CIS ELSA site 2: r = 0.998; IMx vs ES 300: r = 0.980; IMx vs ETI Sorin: r = 0.931. Slopes of regression lines varied from 0.666 for IMx vs ETI Sorin to 0.988 for IMx vs CIS ELSA (site 1, where heparin plasma was used instead of serum as recommended by the manufacturer). Slopes and correlation coefficients were found to be only slightly dependent on assay ranges analysed by statistical procedures applied.

Despite good correlations between methods, it is recommended that samples collected in the follow-up of disease and at higher CA 15-3 concentrations are analysed by the same test; a changeover to another test is not encouraged.

### Introduction

CA 15-3, a breast-cancer-associated antigen, is found in the serum of patients with benign lesions and carcinoma of the breast. The antigen is a high-molecular-mass glycoprotein (a MUC1 gene product) belonging to the family of polymorphic epithelial mucins (1) and is defined by epitopes on the CA 15-3 molecule detected by two different monoclonal antibodies, 115D8 and DF3, respectively (2, 3).

In most currently used assays for CA 15-3 in serum, the antibody 115D8 is used as the catcher antibody, while the DF3 antibody serves as tracer molecule.

CA 15-3 assay values are frequently elevated in patients with breast cancer (4). It was concluded that the CA 15-3 assay may be of clinical value for monitoring the response of patients undergoing therapy. Increasing and decreasing values correlated well with disease progression and regression, respectively (5). Additional published data suggested that, in patients at risk for breast cancer recurrence after primary therapy, increasing CA 15-3 values may indeed be indicative of recurrent disease long before it can be detected clinically (5).

CA 15-3 levels appear to be independent of primary tumour histology. Patients with primary breast cancer may

exhibit elevated serum CA 15-3 concentrations, but less frequently than do patients with metastatic disease. In the latter situation, elevations of CA 15-3 serum levels are correlated with tumour burden and stage.

CA 15-3 is not tumour-specific and the antigen can be detected by monoclonal antibodies in normal and in malignant epithelial tissues. Therefore, increased serum levels can be associated with diseases other than breast carcinoma (6). Other non-mammary malignancies in which elevated CA 15-3 assay values have been reported include lung, colon, pancreatic, primary liver, ovarian, cervical, and endometrial carcinoma (6).

CA 15-3 assay values are elevated in some normal individuals, so that this assay is not to be recommended as a screening procedure for cancer in the general population.

It was the dual purpose of this multicentre study,

(i) to evaluate the analytical performance of the recently introduced CA 15-3 immunoassay adapted to the Abbott IMx immunoanalyser and

(ii) to compare the results with those from the test systems in use in the respective laboratories (Boehringer

ES 300, CIS ELSA, and ETI Sorin). In addition, commercially available quality-control sera were compared. Dilution experiments to test the analytical linearity of the test systems were also included.

## Materials and Methods

### Methods and reagents

Sera and heparin plasma were obtained from randomly selected patients. Samples were stored, either at 4 °C to 8 °C for up to 8 days (site 1) or at -20 °C until assayed (sites 2, 3, and 4).

**Tab. 1** Analytical performance characteristics, manufacturers' claims

	Abbott	BMC	ETI Sorin	CIS ELSA
Methodology	IMx (MEIA)	ES 300 (ELISA)	Streptavidin-biotin (ELISA)	IRMA coated tubes
Antibody combination <sup>a</sup>	115D8/DF3	115D8/DF3	115D8/DF3	115D8/DF3
Assay time	48 minutes	3 hours	> 3.5 hours	2 to 2.5 hours
Calibrators (10 <sup>3</sup> U/l)	0/15/60/120/180/250	0/22/43/101/201	0/10/25/50/100/300	0/15/40/80/140/240
Kit controls (10 <sup>3</sup> U/l)	35/150	2/93	18	30
Sensitivity (10 <sup>3</sup> U/l)	< 0.2	0.9	1	0.2
Cut-off (10 <sup>3</sup> U/l) <sup>b</sup>	28	26	21	30
Sample volume (μl)	150	170	25	20
Automation	yes	yes	no	no
Stored curve	yes	yes	no	no
Sample predilution	no	no	yes	yes
Sample pipetting	yes	yes	yes	yes
Reagent preparation	no	yes	yes	yes
Off-line incubation	no	no	yes	yes
Sample replicates	1	1	2	2

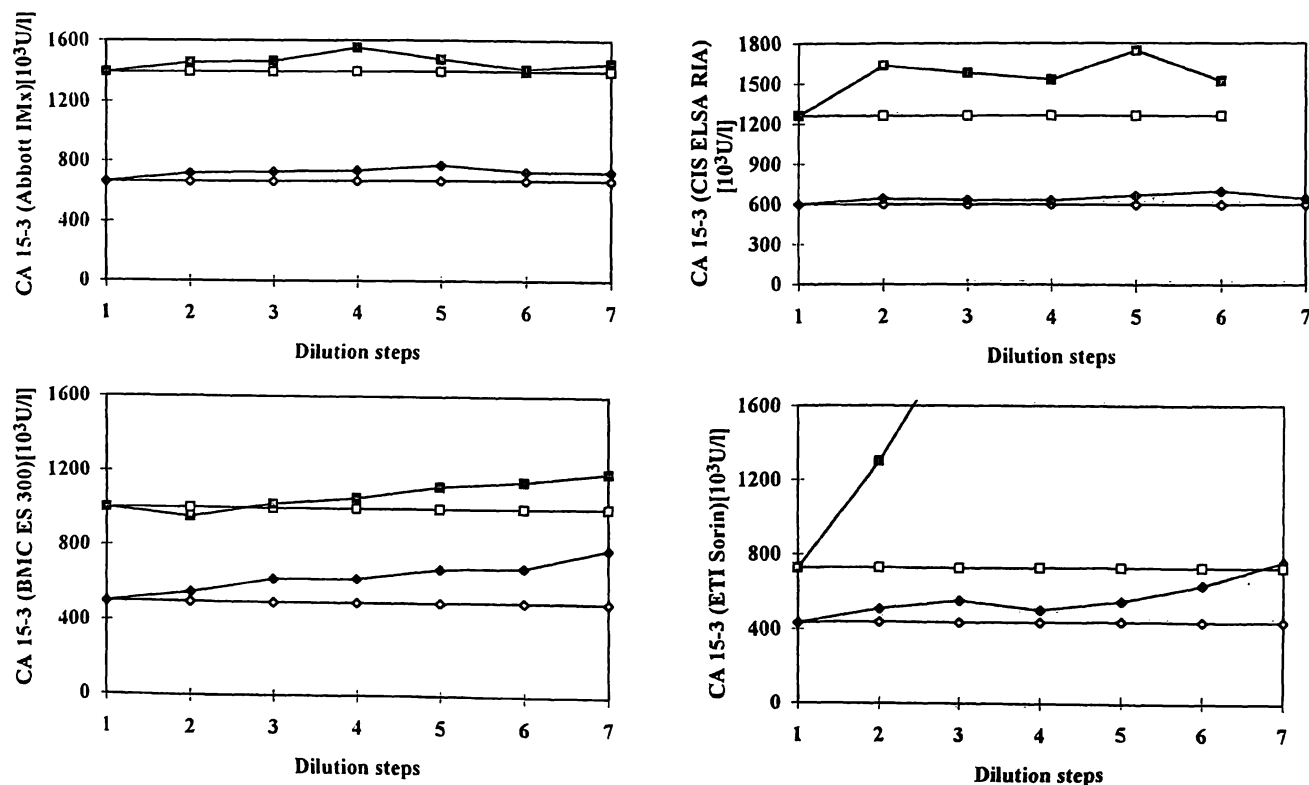
<sup>a</sup> Catcher/tracer monoclonal antibody

<sup>b</sup> Abbott: 95.7% of 648 healthy female subjects had CA 15-3 assay values at or below 28 × 10<sup>3</sup> U/l

Boehringer Mannheim: 95% of 245 healthy subjects with CA 15-3 values between 0–22 × 10<sup>3</sup> U/l, 99% below 30 × 10<sup>3</sup> U/l

ETI Sorin: The observed range for the CA 15-3 concentration in 100 normal sera and normal plasma was 13.0 ± 8 (mean ± SD) × 10<sup>3</sup> U/l

CIS ELSA: In 186 women unaffected by any benign or malignant pathology, 98.6% of the values were below 30 × 10<sup>3</sup> U/l



**Fig. 1** Dilution linearity testing according to the manufacturers' instructions using two patient specimens per test under investigation.

High starting concentration

—□— IMx target —■— IMx analysed  
 —□— ES 300 target —■— ES 300 analysed  
 —□— CIS target —■— CIS analysed  
 —□— Sorin target —■— Sorin analysed

Low starting concentration

—◇— IMx target —◆— IMx analysed  
 —◇— ES 300 target —◆— ES 300 analysed  
 —◇— CIS target —◆— CIS analysed  
 —◇— Sorin target —◆— Sorin analysed

BMC = Boehringer Mannheim Corporation

To analyse CA 15-3, all participating laboratories used the IMx CA 15-3 microparticle enzyme immunoassay (MEIA, lot No. 90191F9) from Abbott AG, Diagnostics Division, Cham, Switzerland. This assay uses monoclonal antibody 115D8-coated microparticles and monoclonal antibody DF3 conjugated to alkaline phosphatase, and is based on IMx MEIA technology (7). Results are based on a stored six-point calibration curve. The IMx assay utilizes a four-parameter logistic curve fit (4PL Analysis) to generate a calibration curve. Assay values are expressed as units per litre (U/l), or  $10^3$  U/l and are related to a reference preparation maintained by Abbott. There is no internationally recognized standard available at this time. Test performances are listed in table 1.

Two of the four participating sites (sites 1 and 2) compared IMx results with those generated with the ELSA CA 15-3 test (lot No. 231A) from CIS BioInternational, Gif-Sur-Yvette, France. This immunoradiometric assay is based on a solid-phase sandwich technique, which also uses the two monoclonal antibodies 115D8 and DF3 described above. The first antibody is coated onto the ELSA solid-phase, the second is radiolabelled with  $^{125}\text{I}$  and used as a tracer. Point-to-point linear interpolation is used to generate the standard curve (8). Predilution of the samples (1/50) with assay buffer is required prior to assay.

At site 3, IMx assay results were compared with those generated by the Enzymun Test CA 15-3 on the ES 300 from Boehringer Mannheim, Rotkreuz, Switzerland (reagent lot No. 645 938 01). This test system consists of a two-step ELISA using streptavidin-biotin technology (9). Results are read from a five-point calibration curve.

Site 4 compared Abbott IMx values with those established with the manual assay ETI-CA-15-3 K from Sorin Biomedica, Antony, France (lot No. 9150010 A and B). Like the other assays under investigation, it utilizes the monoclonal antibodies 115D8 and DF3 and is a one-step sandwich immunoenzymatic assay using streptavidin-biotin technology. The capture antibody (115D8) is conjugated to biotin while the other (DF3) is labelled with horseradish peroxidase. Results are based on a five-point calibration curve and point-to-point linear interpolation (10).

The following commercially available control materials were used: Lyphochek Tumour Marker Control Levels 1 and 2 (BIO-RAD Laboratories AG, Glatbrugg, Switzerland, lot No. 57900 and 64200) and BioRef (BioRef GmbH, Mömbris, Germany) 1 and 2 (lot No. 920 161).

### Imprecision

In order to assess the *intra-assay* imprecision, the internal control solutions from the respective CA 15-3 test systems, Lyphochek tumour marker control levels 1 and 2, the tumour marker control BioRef levels 1 and 2, and a pool of human serum were each measured in 20-fold replicates.

In order to calculate *inter-assay* imprecision, duplicate determinations with IMx and CIS controls as well as Lyphochek levels 1 and 2 or BioRef 1 and 2 were carried out on 10 working days.

### Lower limit of detection

The potential lower limit of detection was determined according to "Clinical Laboratory Improvement Amendment" CLIA (11). Lower limit of detection is defined as the lowest concentration of analyte that can be distinguished from zero with a 95% confidence limit of detection. Following these rules, the zero-calibrators of the individual tests were analysed in replicates of 10 on two different days.

### Dilution linearity

Each laboratory carried out linearity tests. Specimens containing out-of-range elevated levels of CA 15-3 were diluted according to the manufacturer's instructions to provide four or more diluted samples reading in the calibration curve range. All dilutions were run in duplicate in all tests. Linear regression analysis for each of the chosen specimens was performed and the end concentrations calculated (*target values*; cf. fig. 1).

### Quality control of analytical procedures

The test performance was checked by comparing the assigned values of at least three commercial quality control sera with the values obtained for these sera in each run.

### Sample comparison

Patient samples were run in duplicate in the manual methods and singly in the automated methods, and results compared. In each laboratory the method in use was compared with the IMx method. Statistical analysis of results was performed by the method of *Passing & Bablok* (12).

## Results

### Imprecision

#### *Intra-assay variation* (tab. 2):

Pools of locally available sera containing average CA 15-3 concentrations of 10 to  $20 \times 10^3$  U/l and IMx control sera showed coefficients of variation (CVs) well below 5% with the IMx method.

In comparison, the CIS ELSA CA 15-3 test showed CVs between approximately 4 and 9%. With the ETI Sorin test the CVs were well above 10%.

#### *Inter-assay variation* (tab. 2):

All four sites performed tests on the two IMx controls and included the external control sera Lyphochek. As expected, the CVs for the kit's internal and the external control sera were slightly higher than the intra-assay figures, but they stayed within an acceptable range ( $\leq 8\%$ )

**Tab. 2** Imprecision testing

Method/site	Control sample				
	IMx low <sup>1</sup>	IMx high <sup>1</sup>	Pool <sup>3</sup>	Lypho 1 <sup>1</sup>	Lypho 2 <sup>2</sup>
<i>Intra-assay variation</i> <sup>2</sup> (N = 20)					
IMx Site 1	2.2	3.0	(11.8)	2.7	
Site 2	3.4	4.5	(14.2)	2.6	
Site 3	3.9	3.2	(23.4)	2.1	
Site 4	3.0	3.4	(14.0)	1.9	
CIS Site 1	4.2	4.3	(10.5)	5.2	
Site 2			(20.0)	8.6	3.8
ES 300			(13.0)	2.0	2.9
ETI Sorin			(32.2)	22.8	18.7
<i>Inter-assay variation</i> <sup>2</sup> (N = 10)					
IMx Site 1	4.0	5.5		4.2	3.6
Site 2	5.1	7.7		4.5	5.1
Site 3	4.1	6.3		3.8	5.0
Site 4	2.9	7.1		3.5	3.9
CIS Site 1				5.7	2.1
Site 2				19.1	18.8
ES 300				5.3	9.4
ETI Sorin				26.9	11.3

<sup>1</sup> cf. table 3 for mean concentrations

<sup>2</sup> Coefficients of variation (CV %)

<sup>3</sup> Pool: Site-specific pool of sera with an average concentration of CA 15-3 ( $\times 10^3$  U/l)

for the IMx, but not for the CIS ELSA at site 2, or for the ETI Sorin test.

#### Lower limit of detection

The lower limits of detection of the tests [sensitivity ( $10^3$  U/l): mean of zero standards plus 2 standard deviations] claimed by the manufacturers were  $0.2 \times 10^3$  U/l (IMx),  $0.9 \times 10^3$  U/l (ES 300),  $0.2 \times 10^3$  U/l (CIS ELSA) and  $1.0 \times 10^3$  U/l (ETI Sorin). The results obtained in the four laboratories were 0.1, 0.13, 0.07,  $0.06 \times 10^3$  U/l for the IMx, 1.9 and  $0.83 \times 10^3$  U/l for CIS ELSA (site 1 and site 2),  $0.15 \times 10^3$  U/l for the ES 300, and  $3.87 \times 10^3$  U/l for ETI Sorin.

#### Dilution linearity

Figure 1 shows the dilution linearity of the individual tests at two different concentrations levels. It should be noted that the results shown are from four different samples and are a representation of a set of samples run in each laboratory.

The IMx target and analysed values show the closest agreement at levels of approximately 800 and  $1300 \times 10^3$  U/l. In the ES 300 method a tendency to higher values with increasing dilution is observed. Interestingly, the CIS ELSA IRMA test shows good linearity at  $600 \times 10^3$  U/l but varying results at higher concentra-

tions. The ETI Sorin test deviated drastically from the target values.

#### Comparison of analytical procedures

Table 3 summarizes the results of the CA 15-3 determinations in quality control samples measured on different days and compares these values with the assigned concentration ranges for CA 15-3 in these test samples. All the test systems measured CA 15-3 in the targeted concentration ranges. However, the column "Min-Max" (tab. 3) indicates that the concentration ranges are very narrow with the IMx but somewhat broader with the other methods. The ETI Sorin method in particular shows a very broad concentration range. Interestingly, at one site where the BioRef samples were measured by the IMx and by the CIS ELSA, values appeared in rather narrow ranges for IMx and closer to the assigned values for CIS ELSA than CIS ELSA values themselves (tab. 3).

#### Patient samples

Each laboratory assayed approximately 100 serum or plasma samples simultaneously in the routine and IMx assay. Table 4 and figures 2 to 5 show the correlation coefficients and the regression coefficients obtained as analysed using the *Passing-Pablok* method (12) over the whole concentration ranges. These data show that even

Tab. 3 Comparison of mean values (different days) with assigned values of commercially available control samples for all instruments

Method, site	N	Mean ( $\times 10^3$ U/l)	Min-Max ( $\times 10^3$ U/l)	Assigned mean (Assigned range) ( $\times 10^3$ U/l)
<i>Lyphochek 1</i>				
IMx site 1	9	13.0	11.9– 13.9	
IMx site 2	7	12.0	11.3– 12.9	
IMx site 3	10	12.2	11.3– 13.1	
IMx site 4	10	11.3	10.7– 12.0	
ES 300	13	11.4	10.5– 12.3	10.7 (8.6–12.8)
ETI Sorin	10	17.0	13.4– 27.5	13.1 (9.1–17.1)
CIS ELSA site 2	9	13.9	11.5– 20.0	11.0 (9.0–13.0)
<i>Lyphochek 2</i>				
IMx site 1	10	35.7	32.4– 36.6	
IMx site 2	8	34.0	30.4– 36.9	
IMx site 3	10	33.9	31.4– 37.2	
IMx site 4	10	34.2	32.3– 36.5	
ES 300	13	36.8	27.1– 41.4	36.6 (29.3–43.9)
ETI Sorin	10	44.4	37.1– 52.9	38.8 (27.9–49.7)
CIS ELSA site 2	10	42.4	31.9– 61.5	32.0 (26.0–38.0)
<i>BioRef 1</i>				
IMx	10	19.5	17.5– 20.9	17.0 (10.0–23.0)
CIS ELSA	10	22.4	18.6– 27.0	
<i>BioRef 2</i>				
IMx	10	65.7	59.3– 74.4	68.4 (40.0–95.0)
CIS ELSA	10	94.7	84.5–107.0	

**Tab. 4** Comparison of methods, regression analyses ( $y = ax + b$ ) by the method of *Passing & Bablok* (12) (cf. figs. 2 to 5)

Site	x-axis	y-axis	n	r	a	b	Median (min-max) ( $\times 10^3$ U/l)	Comments
1	CIS ELSA	IMx	122	0.950	0.988	0.358	IMx = 28.2 (4.9–10260)	heparin plasma
			98		0.969	0.601	IMx = 19.9 (4.9–248)	heparin plasma
2	CIS ELSA	IMx	101	0.998	0.782	0.064	IMx = 28.0 (4.6–12960)	serum
			86		0.758	0.219	IMx = 22.5 (4.6–245)	serum
3	ES 300	IMx	129	0.980	0.837	1.374	IMx = 31.3 (6.6–3240)	serum
			105		0.762	2.814	IMx = 23.5 (6.6–247)	serum
4	ETI Sorin	IMx	184	0.931	0.666	-0.820	IMx = 15.6 (4.8–250)	all patients, serum
			57	0.952	0.620	0.565	IMx = 15.5 (4.8–250)	breast cancer
			16	0.975	0.733	-2.168	IMx = 18.6 (8.3–89.3)	lung cancer
			20	0.945	0.577	-1.323	IMx = 16.6 (5.1–31.8)	benign breast diseases
			15	0.984	0.678	-1.453	IMx = 15.0 (6.3–26.9)	colorectal cancer
			13	0.963	0.845	-2.537	IMx = 23.4 (9.3–250)	ovarian and uterine cancer
			30	0.871	0.654	0.364	IMx = 15.2 (5.5–42.4)	general diseases
			30	9.852	0.666	0.621	IMx = 14.4 (5.0–24.0)	healthy females

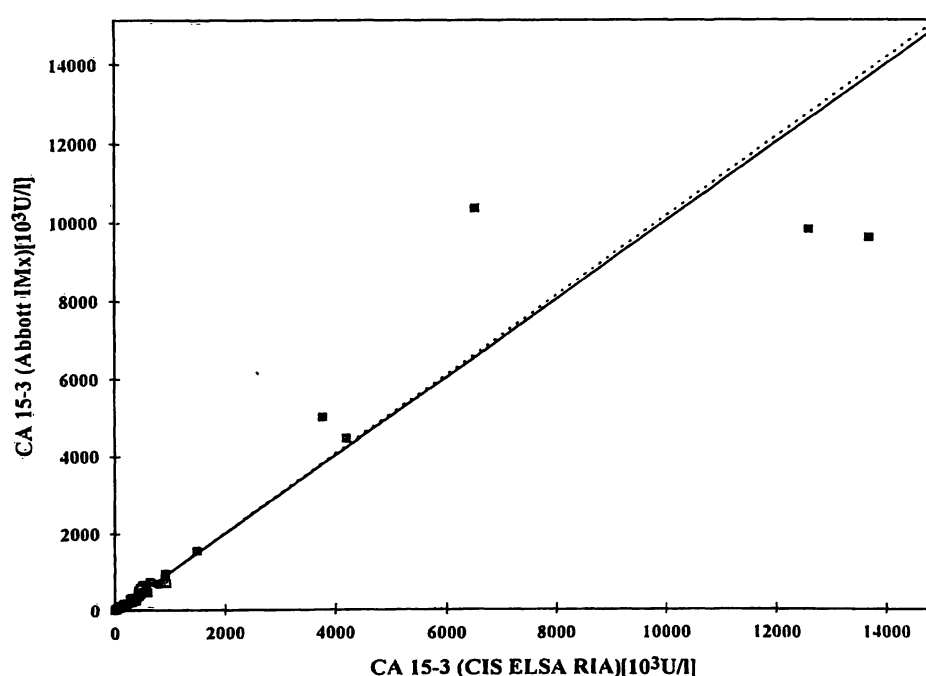
though the correlation coefficients between IMx and CIS ELSA are equal at sites 1 and 2, there is a difference in the slopes of the regression lines (see discussion below). The slope found at site 2 corresponds well to that found at the other two sites where the ES 300 and the ETI Sorin methods were compared with the IMx. At one site (site 4) the test results were broken down into diverse types of non-breast cancers and benign breast lesions; the regression coefficients and the slopes of the regression lines did not vary dramatically from the main population.

As demonstrated in figures 2 to 5, most samples contain CA 15-3 in the concentration range of less than  $50 \times 10^3$  U/l. Values equal to or greater than this concen-

tration tend to vary widely. Dilutions in serum with CA 15-3 concentrations  $> 200 \times 10^3$  U/l were not linear with the ETI Sorin method and are therefore not shown in figure 5. The regression lines deviate from the ideal  $45^\circ$  in all methods using serum samples. This stands in contrast to the comparison study using heparin plasma samples (site 1).

## Discussion

CA 15-3 measurements are widely used in the management of breast cancer treatment and the test is very helpful for the early detection of recurrence of the disease. The sensitivity of the test for indicating small changes in tumour growth justifies automation of the method in



**Fig. 2** Comparison IMx versus CIS ELSA for serum: statistical analysis of results was performed by the method of *Passing & Bablok*.  
 $n = 122$ ,  $r = 0.950$ ,  $y = 0.988x + 0.358$

order to obtain test results quickly. These readily available results allow an effective adaptation of therapy to the growth characteristics of the tumour.

Before using a new test system, it is important for a testing laboratory to know how the new test compares with tests already in use. In this study the automated methods of Abbott IMx and Boehringer ES 300 were compared. With the same study protocol, Abbott IMx

was compared with the manual methods CIS ELSA and ETI Sorin.

The test results of this study show that although the original CIS ELSA test performed satisfactorily, the automation of the method indeed improved the performance. The two automated systems (IMx and ES 300) performed equally well when tested for imprecision.

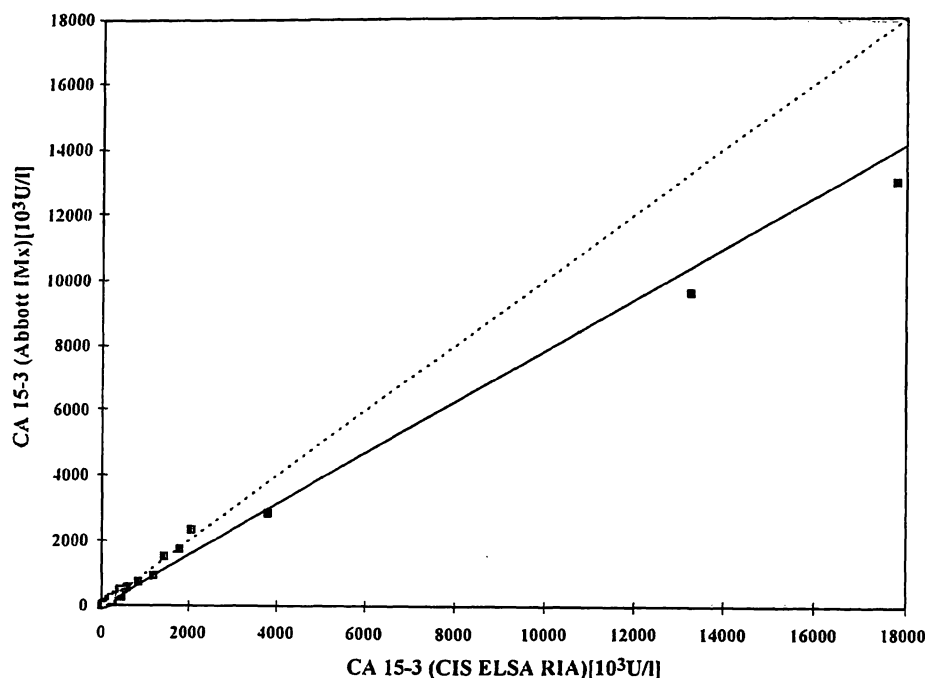


Fig. 3 Comparison IMx versus CIS ELSA for heparin plasma: statistical analysis of results was performed by the method of *Passing & Bablok*.  
 $n = 101$ ,  $r = 0.998$ ,  $y = 0.782x + 0.064$

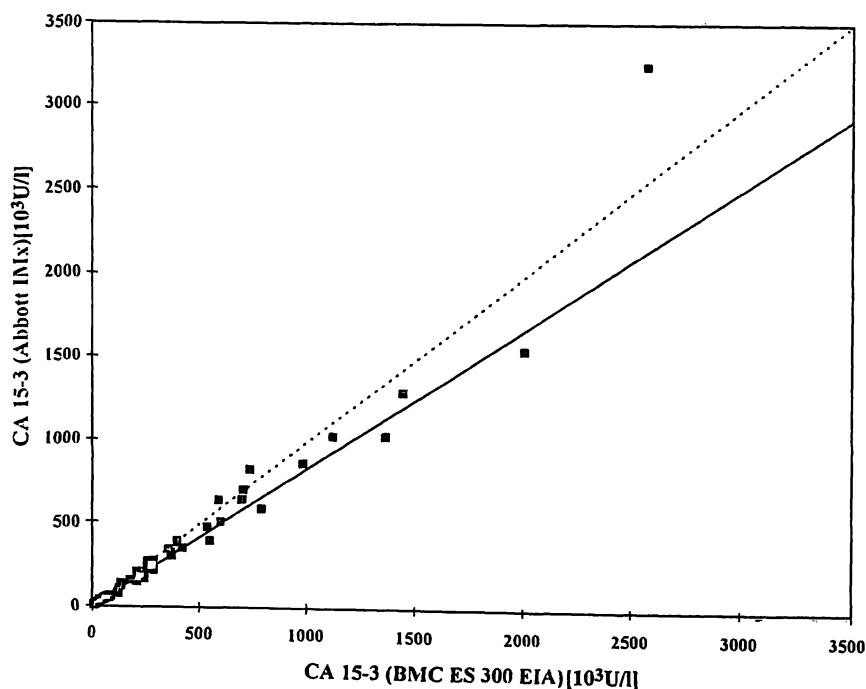


Fig. 4 Comparison IMx versus BMC ES 300 for heparin plasma: statistical analysis of results was performed by the method of *Passing & Bablok*.  
 $n = 129$ ,  $r = 0.980$ ,  $y = 0.837x + 1.374$   
 BMC = Boehringer Mannheim Corporation



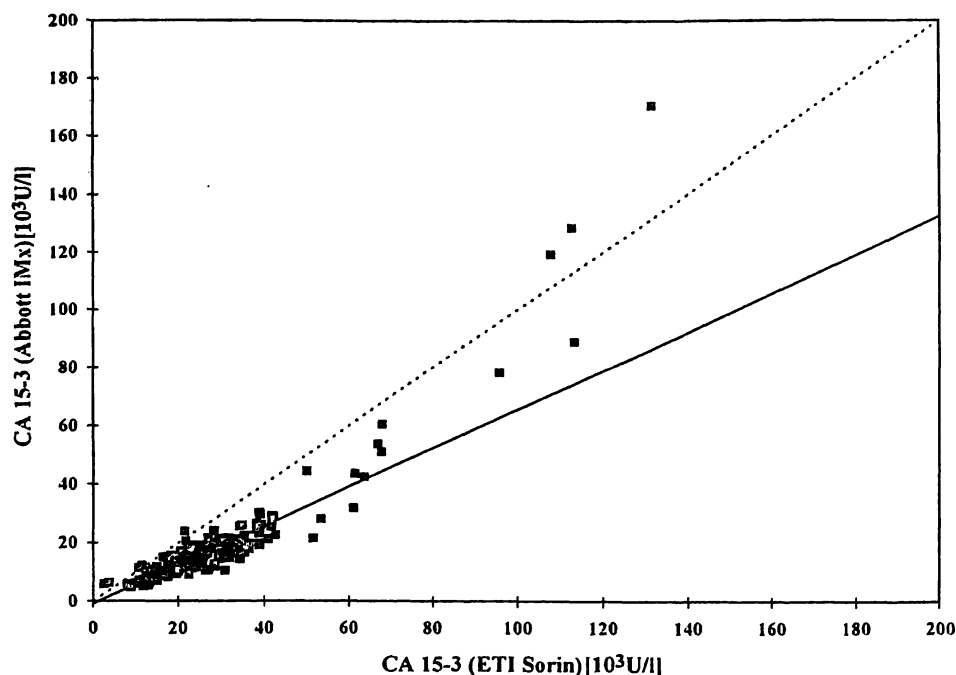


Fig. 5 Comparison IMx versus ETI Sorin: Statistical analysis of results was performed by the method of *Passing & Bablok*.

$n = 181$ ,  $r = 0.931$ ,  $y = 0.666x + 0.820$

As expected and known from experience with manual methods, inter-assay variation was slightly worse than intra-assay variation for the same sample type (tab. 2).

In particular the intra- and inter-assay variations observed with the ETI Sorin test were unsatisfactory. Although the lower limit of detection of the CA 15-3 assay is not a criterion important to the patient, all four methods were sensitive enough and relatively close to the sensitivities claimed by each manufacturer. The linearity tests are of much greater importance, because clinical experience shows that high values of the antigen are to be expected. Therefore, the tests have to yield correct values when samples are diluted serially. Here the automated tests performed well, even though the dilutions had to be carried out manually. The manual tests under investigation seem to be less reliable in this respect and gave dilution values which were too high, reflecting the higher imprecision of these tests (tab. 2).

Day-to-day precision was satisfactory for each test as demonstrated in table 2. Here it is interesting to note that CA 15-3 values obtained by the IMx system with the BioRef quality-control sera, which is apparently produced from cell cultures, are closer to the target values published by the manufacturer for the CIS ELSA method than we found with the same method.

CA 15-3 is recommended as a serum marker for the follow-up of treatment. The necessity to measure accurately high values of the marker must be anticipated. Dilution of the specimen is often necessary.

As one examines the correlation results of the IMx vs the comparison systems under investigation (figs. 2 to

5) the discrepancies between methods became most prominent at higher values, starting at levels between  $2500$  and  $4000 \times 10^3$  U/l for the ES 300 and CIS ELSA (site 1 and 2) assays. Results from the ETI Sorin method start to differ from those from the IMx at much lower concentrations ( $60 \times 10^3$  U/l). In the latter situation severe misinterpretation would occur if methods were changed during the follow-up of disease.

Comparison of patient sample values from the IMx method with those from its competitors shows that the correlation between IMx and CIS ELSA is better when measurements are performed on heparin plasma, rather than on serum. Furthermore, the comparison between IMx and the two other methods (ES 300 and ETI Sorin), also performed on serum samples, gives slopes of the same order as those obtained with the CIS ELSA versus IMx comparison for serum samples. These observations point to a possible effect of heparin; the nature of this effect is unknown, but one might speculate that heparin possibly stabilizes the CA 15-3 molecule by binding to it (CA 15-3 is a mucin and heparin a polyanion; heparin is also used to purify mucin-type proteins). Whatever the mechanism, the practical implication of our observation is that the follow-up of a patient with CA 15-3 measurements should always be performed on the same type of sample and by the same laboratory.

### Acknowledgements

We would like to thank C. Cruffert (Geneva), S. Sander (Basle), and the team of the RIA laboratory (Bern) for their careful work in carrying out the analyses.

## References

1. Hilken J, Buijs F, Hilgers J, Hageman Ph, Calafat J, Sonnenberg A, van der Valk M. Monoclonal antibodies against human milk-fat globule membranes detecting differentiation antigens of the mammary gland and its tumors. *Int J Cancer* 1984; 34:197-206.
2. Hayes DF, Zurawski V, Kufe D. Comparison of circulating CA 15-3 and carcinoembryonic antigen levels in patients with breast cancer. *J Clin Onc* 1986; 4:1542-50.
3. Tobias R, Rothwell C, Wagner J, Green A, Lin Y-SV. Development and evaluation of a radioimmunoassay for the detection of a monoclonal antibody defined breast tumour associated antigen 115D8/DF3. *Clin Chem* 1985; 31:986.
4. Colomer R, Ruibal A, Genolla J, Rubio D, Del Campo JM, Bodi R, Salvador L. Circulating CA 15-3 levels in the postsurgical follow-up of breast cancer patients and in non-malignant diseases. *Breast Cancer Res Treat* 1985; 13:123-33.
5. Safi F, Kohler I, Röttinger E, Beger HG. The value of the tumour marker CA 15-3 in diagnosing and monitoring breast cancer. *Cancer* 1991; 68:574-82.
6. Colomer R, Ruibal A, Genolla J, Salvador L. Circulating CA 15-3 antigen levels in nonmammary malignancies. *Br J Cancer* 1989; 64:1674-81.
7. Fiore MD, Mitchell JE, Doan T, Nelson R, Winter G, Grandone C, et al. The Abbott IMx automated benchtop immunochemistry analyzer system. *Clin Chem* 1988; 34 (9):1726-32.
8. Ruibal A, Genolla J, Rosell M, Moragas M. El CA 15-3 ELSA serio en patalogias no tumorales [letter]. *Med Clin (Barc)* 1987; 88:476.
9. Package insert: BMC ES 300 and Enzymun CA 15-3 Test.
10. Package insert: Sorin ETI-CA-15-3-Test.
11. CLIA 88: Clinical Laboratory Improvement Amendments of 1988, final rule, Federal Register / Vol. 57, No. 40, 7001-7288 / Part II (1992), United States Government Printing Office.
12. Passing H, Bablok W. A new biometrical procedure for testing the equality of measurements from two different analytical methods: application of linear regression procedure for method comparison studies in clinical chemistry. Part I. *J Clin Chem Clin Biochem* 1983; 21:709-20.

*Received May 29/September 11, 1995*

Corresponding author: PD Dr. R. Peter Huber, Hormonlabor, DZL, Kantonsspital Basel, CH-4031 Basel, Switzerland